A Two-Sided Discussion of Preregistration of NLP Research

Anders Søgaard,
Daniel Hershcovich,
Miryam de Lhoneux

EACL 2023

UNIVERSITY OF COPENHAGEN

KU LEUVEN

Preregistration

*Preregistration refers to the practice of specifying what you are going to **do**, and what you expect to **find** in your study, **before** carrying out the study.*

Process:

1. Register hypothesis, experimental design and analysis
2. Submit time-stamped private/public preregistration
3. Follow the plan as closely as possible
4. Submit final study for publication

Goal:
- Reproducible science
- Distinguish **exploratory** and **confirmatory** research

Preregistering NLP research (van Miltenburg et al., NAACL 2021)

Daniel: Anders, there's an additional argument for preregistration, I believe. Early feedback on experimental methodology through a peer-reviewed registration process should improve the quality of the methodology, should it not? Such feedback also saves resources otherwise spent on failed or misleading experiments.

Daniel: NLP as a field has many virtues, but the reviewing cycle is slowing as the field grows larger. Moreover, experiments are becoming more expensive with larger models, creating barriers of entry and experiments have substantial environmental impact.

Anders: Daniel, we both know turn-around is fast in NLP research. Experiments are easier to run and feedback is much faster than for clinical trials, where preregistration is common.

Anders: You make an important point, Daniel, but early feedback would require more time from reviewers. Since reviewers and researchers coincide, preregistration would potentially save compute resources, but not working hours.

Read the paper for more dialogue on why...

✓ Preregistration may bias NLP toward confirmatory research

✓ Preregistration should allow for re-classification of confirmatory research as exploratory research

✓ Preregistration may increase publication bias

✓ Preregistration may increase flag-planting

✓ Preregistration may increase p-hacking

✓ Preregistration may lower our risk tolerance

And for a practical proposal for implementing preregistration in NLP research